Eshan Bhimani

U.S. Citizen | (404)-259-5936 | bhimanieshan@gmail.com | linkedin.com/in/eshan-bhimani | eshan-website-tau.vercel.app

Education

Georgia Institute of Technology — Atlanta, GA
Bachelor of Science in Computer Science, Threads: Systems & Architecture, Intelligence — *May 2028*
GPA: 3.94
Relevant Courses: Data Structures, Algorithms, Computer Organization, AI, Machine Learning, Database
Honors: Presidential Scholar, Zell Miller Scholar, Dean's List

Experience

NCR Atleos — May 2026 – August 2026
Software Engineer Intern — *Atlanta, GA*

- Built a Graph RAG query system for the *Data & AI Team* on Microsoft Fabric using LangChain, Neo4j, and LLM-generated Cypher queries, reducing agent response latency *70%* and improving answer accuracy *50%*
- Developed an agentic schema-recommendation engine with Azure OpenAI that profiled raw SQL tables and generated Power BI-ready semantic models, reducing setup time *60%* and automating schema scaffolding
- Architected a Neo4j-backed data layer integrating 5+ internal data sources and 10M+ relationships into a unified graph backend, reducing multi-hop retrieval latency *50%*

Convention Connection — December 2025 – July 2026
Software Engineer Intern — *Atlanta, GA*

- Engineered an AI-powered image processing pipeline using OpenCV and FastAPI to automate card photo cropping, reducing manual processing time by *75%* and managing *1000+* images via Google Photos API on GCP
- Built a Python pricing engine with real-time marketplace bid monitoring, automating valuation across inventory
- Built a trade matching engine with weighted scoring across grade, set, and company to rank comparable sales

Algoverse AI — September 2025 – March 2026
AI Researcher — *Palo Alto, CA*

- Developed a novel multi-task benchmark evaluating LLMs' vision-based capabilities on geolocation reasoning and tool use; designed spatial reasoning evaluations and submitted the work to *COLM 2026*
- Automated collection and validation of *300+* geolocation samples using Python, Selenium, and the GSV API
- Productionized the research codebase with CI/CD and 50+ automated tests and type checks, reducing setup time *20%* and enabling one-command reproducible evaluation

Projects

Polymarket HFT Bot: Low Latency Systems | *C++23, Python, PyTorch, pybind11, Goldsky* — April 2026

- Engineered a high-frequency trading framework in C++ and Python targeting Polymarket BTC 15-minute prediction markets, integrating Polymarket CLOB API and Goldsky for real-time on-chain market data ingestion
- Implemented cross-venue arbitrage detection across Polymarket, Kalshi, and Binance using a Log-Normal probability model with a PPO reinforcement learning agent and LSTM actor-critic network built in PyTorch
- Bridged C++23 execution engine with Python ML inference via pybind11 for *sub-millisecond* order routing

Switch: AI Model Routing Platform | *TypeScript, Next.js, pgvector, LangGraph* — February 2026

- Launched an AI model-router that directs prompts to the best-fit model for each task, serving *1,000+* users
- Built and fine-tuned a prompt-routing classifier that split traffic 70/30 between economy and premium models; unified 4 provider APIs behind middleware for consistent streaming, model fallback, and request handling
- Engineered Cortex, a persistent memory system using Supabase PostgreSQL and pgvector, enabling low-latency semantic retrieval of user context and conversation history supporting *13k+* daily database requests across LLMs

Technical Skills

Languages: Python, Java, C++, TypeScript, JavaScript, SQL, Linux
AI Tools: LangChain, LangGraph, Agno Agent, Claude Code, Azure OpenAI, pgvector, MCP, Antigravity, Copilot
Frameworks: React, Next.js, FastAPI, Django, PyTorch, Node.js, D3.js, Spring Boot
Tools & Libraries: Git, Docker, CI/CD, AWS, Vercel, Neo4j, PostgreSQL, Supabase, Redis, OpenCV, pybind11